

## **GENE MARKERS FOR CHRONIC MUCOSAL INJURY**

### **BACKGROUND OF THE INVENTION**

Clinical assessment of disease activity in ulcerative colitis or Crohn's disease is very difficult. Patient symptoms do not necessarily correlate with the inflammatory (disease) activity in the small intestine and colon, leading to educated guesses being used to direct anti-inflammatory therapy. Similar difficulty exists in measuring or testing the efficacy of new therapeutic compounds. Currently the gold standard in diagnosing ulcerative colitis or Crohn's disease is the use of fiberoptic endoscopy coupled with multiple biopsies and pathologic analysis. This very expensive approach requires a skilled specialist and has associated risks, such as risk of sedation, bleeding, and colon perforation. The patient is also subjected to discomfort from the procedure and preparation.

A less invasive and less risky assessment of mucosal disease activity is needed to accurately guide treatment and to provide an objective measure of mucosal injury, both for patients and for use in clinical studies. There is also a need for a simple test to aid in the differentiation of chronic inflammatory disease (UC or CD) from common acute inflammatory disorders or common non-inflammatory benign disorders. There is a further need for a simple method for the differentiation of ulcerative colitis and Crohn's disease because the surgical and

medical management for these two diseases is profoundly different.

## **SUMMARY OF THE INVENTION**

It is an object of the invention to provide a method for identifying chronic mucosal injury in a human.

It is another object of the invention to provide a method of determining the degree of injury to the small intestine or colon of a human with chronic mucosal injury.

It is yet another object of the invention to provide a method for monitoring the efficacy of therapy for chronic mucosal injury.

It is a further object of the invention to provide a method of screening compounds for anti-chronic mucosal injury or anti-ulcerative colitis activity.

These and other objects of the invention are provided by one or more of the embodiments described below.

One embodiment of the invention provides a method of diagnosing chronic inflammatory bowel disease. At least one gene expression product of the *regenerating (REG)* gene family is detected in a body sample of a human who is suspected of having chronic inflammatory bowel disease. The human is identified as having chronic inflammatory bowel disease if the gene expression product is detected.

A further embodiment of the invention provides a method to aid in the differentiation of chronic mucosal injury from common acute inflammatory colon disorder and common non-inflammatory benign colon disorder in a human with symptoms of bowel disease. The amount of at least one gene expression product of the *REG* gene family in a body sample of a first human who is suspected of having bowel disease, is compared with the amount of the gene expression

product in a body sample of a second human who is healthy. The first human is identified as having chronic mucosal injury if the body sample of the first human contains more of the gene expression product than the body sample of the second human.

Another embodiment of the invention provides a method to determine degree of injury to small intestine or colon tissue of a human with chronic mucosal injury. A quantity of a gene expression product of the *REG* gene family in a body sample of a human having chronic mucosal injury is determined. The amount is correlated with the degree of injury to the small intestine or colon.

Still another embodiment of the invention provides a method of monitoring the efficacy of therapy for chronic mucosal injury in a human body sample. At least one gene expression product of the *REG* gene family is quantitated in a body sample of a human who has been subjected to therapy for chronic mucosal injury. The quantity of the expression product in the sample is compared to the quantity of the gene expression product in a matched body sample of the human at an earlier time. A reduction in the quantity of the gene expression product after therapy is an index of efficacy of the therapy.

Another embodiment of the invention provides a method of screening compounds for anti-chronic mucosal injury activity. A colonic cell expressing a gene which is a member of the *REG* gene family is contacted with a test compound. The expression of the *REG* gene is quantitated. A test compound which decreases expression of the gene is identified as a potential compound for treating chronic mucosal injury.

A further embodiment of the invention provides a method of diagnosing ulcerative colitis. An mRNA which is expressed by a gene represented by a Hs.111244 polynucleotide is detected in

a body sample of a first human who is suspected of having ulcerative colitis. The human is identified as having ulcerative colitis if the mRNA is detected.

Still another embodiment of the invention provides a method to aid in the differentiation of ulcerative colitis from common acute inflammatory colon disorder, common non-inflammatory benign colon disorder, and Crohn's disease in a human with symptoms of bowel disease. The amount of mRNA which is expressed by a gene represented by a Hs.111244 polynucleotide in a body sample of a first human suspected of having bowel disease is compared with the amount of the mRNA in a comparable body sample of a second human who is healthy. A body sample of the first human which contains more of the mRNA than the body sample of the second human identifies the first human as having ulcerative colitis.

Another embodiment of the invention provides a method to determine the degree of injury to small intestine or colon tissue of a human with ulcerative colitis. A quantity of an mRNA which is expressed by a gene represented by a Hs.111244 polynucleotide in a body sample of a first human having ulcerative colitis is determined. The quantity of the mRNA is correlated with the degree of injury to the small intestine or colon.

Even another embodiment of the invention provides a method of monitoring the efficacy of therapy for ulcerative colitis in a human body sample. An mRNA which is expressed by a gene represented by a Hs.111244 polynucleotide is quantitated in a body sample of a human who has been subjected to therapy for ulcerative colitis. The quantity of the mRNA in the sample is compared to the quantity of the mRNA in a matched body sample of the human at an earlier time. A reduction in the quantity of the mRNA after therapy is an index of efficacy of the therapy.

Still another embodiment of the invention provides a method of screening compounds for

anti-ulcerative colitis activity. A colonic cell expressing an mRNA which is expressed by a gene represented by a Hs.111244 polynucleotide is contacted with a test compound. The expression of the mRNA by the cell is quantitated. A test compound which decreases expression of the mRNA is identified as a potential compound for treating ulcerative colitis.

#### **BRIEF DESCRIPTION OF THE DRAWINGS**

Figure 1 demonstrates the hybridization signal of spiked controls. The fluorescence intensity for different levels of gene expression was standardized by spiking a known amount of control genes.

Figure 2 demonstrates the expression of *PSP*, *PAP*, and *REGH* in inflammatory bowel disease.

Figure 3 demonstrates the use of reverse transcriptase PCR with primers specific for *PAP*, *PSP*, or *REGH* on mRNA isolated from a healthy human and a human with ulcerative colitis.

Figure 4 demonstrates the purification of recombinant *PSP* and *REGH* proteins.

#### **DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS**

The inventor has discovered that chronic mucosal injury can be diagnosed by detecting expression levels of the *REG* gene family and a gene represented by a Hs.111244 polynucleotide in a human body sample. The members of the *REG* gene family have been found to be strongly expressed in regions of the colon involved with chronic mucosal injury and in the small intestine and colon of humans with Crohn's disease. Additionally, a gene represented by the Hs.111244 polynucleotide is strongly expressed in the colon of humans with ulcerative colitis. Such strong

expression is both surprising and useful because the *REG* gene family and the gene represented by the Hs.111244 polynucleotide are not expressed, or are expressed at low levels, in healthy small intestine and colon tissue. Further, the expression products of the *REG* gene family and the gene represented by the Hs.111244 polynucleotide can be detected in the serum of humans with chronic mucosal injury and ulcerative colitis, respectively.

Chronic mucosal injury can be caused by inflammatory bowel diseases such as ulcerative colitis and Crohn's disease. Crohn's disease affects both the small intestine and the colon. Chronic mucosal injury can further be caused by immunodeficiencies, such as chronic granulomatous disease and transplantation rejection, and infections, such as mycobacteria.

At the present time, the human regenerating (*REG*) gene family is known to contain four genes: *pancreatic stone protein (PSP)* (the protein is also known as thread protein, lithostathine, and Reg) as shown in SEQ ID NO:1, *pancreatitis-associated protein (PAP)* as shown in SEQ ID NO:2, *human pancreatic beta cell growth factor*, also known as (*INGAP*), as shown in SEQ ID NO:3, and *regenerating gene homologue (REGH)* as shown in SEQ ID NO:4. In normal circumstances, these genes are regionally expressed in low amounts in the small bowel and pancreatic epithelium. Healthy colonic mucosa and small intestine has little or no expression of the *REG* gene family. Ulcerative colitis, Crohn's disease, or other chronic mucosal injury leads to high levels of *REG* gene expression in the colonic mucosal or small intestine or both. This expression of the members of the *REG* gene family correlates with the degree of histopathological injury, and is not seen in the setting of acute self-limited colonic inflammation or common non-inflammatory benign colon disorders.

The nucleic acid sequence of the expressed sequence tag (EST) Hs.111244 is shown in

SEQ ID NO:5. The nucleic acid sequence of Hs.111244 has been newly determined and represents a more complete sequence of Hs.111244 than has been previously published. This nucleic acid sequence is referred to herein as the Hs.111244 polynucleotide. The gene represented by the expressed sequence tag (EST) Hs.111244, is not expressed or is expressed at low levels in the healthy colon mucosa, acute self-limited colonic inflammation or common non-inflammatory benign colon disorders, or mucosa affected by Crohn's disease. However, the gene represented by the Hs.111244 polynucleotide is expressed at high levels in colonic mucosa affected by ulcerative colitis.

In humans who have been diagnosed with a bowel disease, detection of levels of at least one gene expression product of the *REG* gene family in a body sample can be used to diagnose or prognose chronic mucosal injury or to monitor treatment of chronic mucosal injury. The body sample is obtained from a human and can be, for example, a tumor, a solid tissue such as colon or small intestine tissue, or a fluid sample such as blood, serum, or plasma. The human from whom the body sample is obtained can be apparently healthy or can already be identified as having chronic mucosal injury. A comparable body sample is a body sample obtained from a second human which is the same type of body sample as obtained from a first human. A matched body sample is a body sample obtained from the first human at an earlier time which is the same type of body sample from the first human obtained at a later time.

Expression products of the *REG* gene family can be detected in a body sample. Detection of the expression products in a human's body sample indicates the presence of chronic mucosal injury in the human. In one embodiment, the body sample is assayed for the presence of at least one *REG* gene family protein. A *REG* gene family protein or polypeptide, can be detected using,

for example, anti-REG gene family-specific antibodies. The antibodies can be labeled, for example, with a radioactive, fluorescent, biotinylated, or enzymatic tag and detected directly, or can be detected using indirect immunochemical methods, using a labeled secondary antibody. The presence of REG gene family protein or polypeptides can be assayed, for example, in tissue sections by immunocytochemistry, or in lysates, using Western blotting, as is known in the art. Further, REG gene family proteins or polypeptides can be assayed by immunoprecipitation assay, enzyme-linked immunoabsorbant assay, quantitative antigen capture-based immunoassay, and radioimmunoassay.

The level of at least one REG gene family protein or polypeptide in a body sample of a human suspected of having a chronic mucosal injury can be compared with the level of the protein or polypeptide in a healthy body sample. The level of a REG gene family protein or polypeptide in a body sample of a human suspected of having chronic mucosal injury can be determined using antibodies specific for the REG gene family protein or polypeptide. The level of the REG gene family protein or polypeptide in a healthy body sample can also be determined. The two levels are compared to each other and a higher level of the REG gene family protein or polypeptide in the suspect human's body sample as compared to the healthy human's body sample indicates the presence of chronic mucosal injury in the suspect human. Preferably, the increased level of the REG gene family protein in the suspect sample is at least 25%, 50%, 100%, 150%, 200% or 250% higher than in the healthy body sample.

Alternatively, the presence of mRNA expressed from at least one member of the REG gene family or mRNA expressed from the gene represented by the Hs.111244 polynucleotide can be detected in a body sample. Detection of mRNA expressed from at least one member of the

*REG* gene family in a body sample of a human indicates the presence of chronic mucosal injury in the human. Detection of mRNA which is expressed by a Hs.111244 polynucleotide in a body sample of a human suspected of having bowel disease indicates the presence of ulcerative colitis in the human.

mRNA expressed from the *REG* gene family or the gene represented by a Hs.111244 polynucleotide can be detected by any means known in the art. For example, one can use *in situ* hybridization in tissue sections or Northern blots containing poly A<sup>+</sup> mRNA. Other techniques such as high density DNA array hybridization, ribonuclease protection assay, and serial analysis of gene expression can also be used. *REG* gene family- or Hs.111244-specific oligonucleotide probes can be generated using the polynucleotide sequences of the *REG* gene family or of the gene represented by a Hs.111244 polynucleotide. The probes are preferably at least 12, 14, 16, 18, 20, 22, 24, or 25 nucleotides in length and can be less than 2, 1, 0.5, 0.1, or 0.05 kb in length. The probes, for example, can be synthesized chemically, generated from longer polynucleotides using restriction enzymes, or amplified enzymatically. The probes can be labeled, for example, with a radioactive, biotinylated, or fluorescent tag. A mixture of probes can also be used. Such a mixture can contain a plurality of probes which are specific to different *REG* family genes or specific for the gene represented by the Hs.111244 polynucleotide. Alternatively, each of a plurality of probes can be used separately.

One of skill in the art can readily determine differences in the amount of *REG* gene family mRNA or a gene represented by a Hs.111244 polynucleotide mRNA transcripts between two body samples, for example, using Northern blots and nucleotide probes. The level of mRNA expressed from of at least one member of the *REG* gene family or the gene represented by the

Hs.111244 polynucleotide in a body sample of a human suspected of having chronic mucosal injury, can be compared with the mRNA expression from at least one member of the *REG* gene family or the gene represented by Hs.111244 polynucleotide in a healthy body sample. This can be done, for example, using *in situ* hybridization in tissue section or in Northern blots containing poly A<sup>+</sup> mRNA. A higher level of mRNA expressed from a gene represented by a Hs.111244 polynucleotide in the suspect body sample as compared to the healthy body sample is indicative of ulcerative colitis in the suspect human who has provided the body sample. A higher level of mRNA expressed from a *REG* family gene in the suspect body sample as compared to the healthy body sample is indicative of chronic mucosal injury in the suspect human who has provided the body sample. Preferably, the increased level of mRNA expressed from a member of the *REG* gene family or the gene represented by the Hs.111244 polynucleotide in the suspect body sample is at least 25%, 50%, 100%, 150%, 200%, or 250% higher than in the healthy body sample.

If desired, the level of a particular mRNA, polypeptide, or protein expressed from a *REG* gene family member or mRNA expressed from a gene represented by a Hs.111244 polynucleotide in a body sample can be quantitated. Quantitation can be accomplished, for example, by comparing the level of expression product detected in the body sample with the level of expression product present in a standard curve. A comparison can be made visually or using a technique such as densitometry, with or without computerized assistance.

In a preferred embodiment, chronic mucosal injury can be differentiated from common acute inflammatory colon disorder and common non-inflammatory benign colon disorder in a human with symptoms of bowel disease. The amount of at least one gene expression product such as mRNA or protein of the *REG* gene family in the suspect body sample is compared to the

amount of the same gene expression product in a body sample of a human which is healthy. The gene expression products in the two samples can be compared by any means known in the art. A body sample from a human suspected of having bowel disease which contains more of the gene expression product than the body sample of the healthy human identifies the suspect human as having chronic mucosal injury. Preferably the amount of the gene expression product in the body sample of the human with chronic mucosal injury is increased by at least 25%, 50%, 75%, 100%, 200%, or 250%.

Further, ulcerative colitis can be differentiated from common acute inflammatory colon disorder, common non-inflammatory benign colon disorder, and Crohn's disease in a human with symptoms of bowel disease. The amount of mRNA which is expressed by a gene represented by a Hs.111244 polynucleotide in a first body sample of a human suspected of having bowel disease is compared with the amount of the mRNA in a body sample of a second human which is healthy. The amount of mRNA in the two samples can be compared by any means known in the art. A body sample from a human suspected of having bowel disease which contains more of the mRNA than the body sample of the healthy human identifies the suspect human as having ulcerative colitis. Preferably the amount of mRNA in the body sample of the human with ulcerative colitis is increased by at least 25%, 50%, 75%, 100%, 200%, or 250%.

The degree of injury to the small intestine or colon tissue of a human with chronic mucosal injury can be determined by measuring the quantity of a gene expression product, such as mRNA or protein, of the *REG* gene family in a body sample of the human. The quantity of the gene expression product is correlated with the degree of injury to the small intestine or colon.

Further the degree of injury to the colon tissue of a human with ulcerative colitis can be

determined by measuring the quantity of a mRNA which is expressed by a Hs.111244 polynucleotide or the gene represented by it in a body sample of the human. The quantity of the mRNA is correlated with the degree of injury to the colon.

The efficacy of therapy for chronic mucosal injury can be monitored in a body sample of a human with chronic mucosal injury. At least one gene expression product of the *REG* gene family can be quantitated in a body sample of a human which has been subjected to therapy for chronic mucosal injury. The quantity of the gene expression product in a matched body sample is compared to the quantity of the gene expression product in the body sample at an earlier time. A reduction of in the quantity of the gene expression product after therapy is an index of efficacy of the therapy. Preferably, the amount of the gene expression product is decreased by at least 10%, 25%, 50%, 75% or 100%.

The efficacy of therapy for ulcerative colitis in a human body sample can also be monitored. An mRNA which is expressed by a Hs.111244 polynucleotide or the gene represented by it is quantitated in a body sample of a human which has been subjected to therapy for ulcerative colitis. The quantity of the mRNA in the sample is compared to the quantity of the mRNA in the matched body sample at an earlier time. A reduction of in the quantity of the mRNA after therapy is an index of efficacy of the therapy. Preferably, the amount of the mRNA is decreased by at least 10%, 25%, 50%, 75% or 100%.

According to another aspect of the invention, compounds which have anti-chronic mucosal injury or anti-ulcerative colitis activity can be identified. A colonic cell expressing a gene of the *REG* family or a gene represented by the Hs.111244 polynucleotide can be contacted with a test compound. The test compound can be a pharmacologic agent already known in the art or can

be a compound previously unknown to have any pharmacological activity. The test compound can be naturally occurring or designed in the laboratory. It can be isolated from microorganisms, animals, or plants, and can be produced recombinantly, or synthesized by chemical methods known in the art.

The cell can be any primary human cell or human cell line which expresses a *REG* family gene, or a gene represented by the Hs.111244 polynucleotide, as disclosed above. Methods of establishing cultures of primary human cells or of culturing cell lines are well known in the art.

Expression of at least one gene of the *REG* gene family or the gene represented by the Hs.111244 polynucleotide can be monitored. Expression can be measured in a sample of the same cell population before and after contact with the test compound. Alternatively, control cell populations can be employed. A test compound which decreases expression of at least one member of the *REG* gene family is identified as a potential drug for chronic mucosal injury. A test compound which decreases expression of a gene represented by the Hs.111244 polynucleotide is identified as a potential drug for decreasing ulcerative colitis. Preferably, the test compound decreases the amount of the gene expression product by at least 10%, 25%, 50%, 75% or 100%.

#### **SEQUENCE LISTING**

**SEQ ID NO:1** cDNA sequence of *pancreatic stone protein (PSP)*

**SEQ ID NO:2** cDNA sequence of *pancreatitis-associated protein (PAP)*

**SEQ ID NO:3** cDNA sequence of *human pancreatic beta cell growth factor (INGAP)*

**SEQ ID NO:4** cDNA sequence of *regenerating gene homologue (REGH)*

**SEQ ID NO:5** cDNA sequence of Hs.11124

**Example 1**

*PSP*, *PAP* and *REGH* are expressed in colonic mucosa of patients with inflammatory bowel disease. Parallel methods of measuring gene expression have been recently developed which allow concurrent measurement of the expression of a large number of genes. Light-directed solid-phase combinatorial chemistry was used to generate oligonucleotide probe arrays which provide representation of nearly 7000 human cDNA and EST sequences. Each gene is represented by 20 individual 25-mer oligonucleotide sequences. mRNA isolated from the mucosa of colonic resection specimens was used to generate hybridization probes for our analysis. Details of the GENECHIP technology, probe synthesis, hybridization, and confocal scanning have been previously described. The fluorescence intensity for different levels of gene expression was standardized by spiking known amounts of control genes into the probe mixture (Figure 1). Detection at 1.5 pM is approximately equal to one message copy per cell. Tissue samples taken from the area used to isolate RNA were sent for histochemistry to be scored for acute and chronic inflammation, ulceration, dysplasia, eosinophilia, epithelial apoptosis, and metaplastic changes. Expression levels of *PSP*, *PAP*, and *REGH* in 15 clinical specimens are shown in Figure 2. *PSP* RNA expression was in the top 2% of all arrayed genes in ulcerative colitis. Expression levels corresponded closely to histologic measures of disease activity. One non-IBD patient with severe acute inflammation, but no ulceration (rectal prolapse-specimen 11), did not express detectable *PSP*, *PAP*, or *REGH*.

### Example 2

*Pichia pastoris* expression vectors were constructed with *PSP*, *PAP*, and *REGH*. Gene-specific primer pairs were designed to incorporate a 5' Xho I site, and a portion of the  $\alpha$ -factor gene leading up to the yeast Ste13 cleavage site (5' end) and a Xba I containing primer that deleted the

stop codon and included a 3' Myc epitope tag. After Kex2 and Ste13 signal cleavage by *Pichia*, the amino-terminus should be identical to the native secreted protein. RT-PCR was performed using the TITAN One Tube System (Boehringer Mannheim). The PCR products of these reactions, using RNA from a healthy or UC patient are shown in Figure 3. These results demonstrate trace amounts of *PAP* in the "normal" patient examined, but otherwise agree with the results of the GENECHIP hybridization. Bands were gel purified and cloned into pGEM-T (Promega). The *Xba/Xho* gene fragments were excised and ligated into pPICZ( (Invitrogen). Constructs were bidirectionally sequenced with primers derived from the vector and were found to match the published sequences. Linearized plasmids were transformed into *Pichia* KM71 and recombinant clones identified by Zeocin selection.

### Example 3

PSP, PAP, and REGH were expressed in *Pichia pastoris*. PICZ $\alpha$  places the inserted gene downstream of a strong methanol-inducible AOX1 promoter. Individual clones were grown in 10 ml cultures of BMGY media overnight and resuspended into 1/5 volume of BMMY (0.5% methanol) for induction. Aliquots of media supernatant were taken at various times after methanol induction and subjected to 15% SDS-PAGE. PSP or REGH expression was identified by the new appearance of 18 and 18.5 kDa bands (respectively), peaking 48 hours after induction. These sizes include the 2.5 kDa C-terminal epitope tag, whose presence was verified by Western blot (ECL) utilizing a monoclonal anti-Myc antibody (Invitrogen). Tryptic digestion of both PSP and REGH led to a mobility reduction of about 1 kDa, reflecting the expected size change following cleavage at the Arg11-Ile12 bond (data not shown). PAP expression was also demonstrated.

5

Large scale protein purification was performed by directly scaling up the protocol outlined above. *Pichia* from 1.5 liter cultures were resuspended into 300 mls of BMMY induction media and allowed to grow for 48 hours. Culture supernatants were concentrated by ammonium sulfate precipitation and column purified by Bio-Gel P30. Fractions containing PSP or REGH were pooled and concentrated using a Centriprep 10 concentrator (Amicon). Figure 4 shows a Coomassie-stained SDS-PAGE gel of the purified proteins.